数理解析研究所講究録
第 1796 巻 2012 年 109-114

有限集団における繰り返し囚人のジレンマゲームの終盤効果の検討

東京工業大学大学院・日本学術振興会　関口　卓也 (Takuya Sekiguchi)
Tokyo Institute of Technology / Japan Society for the Promotion of Science

1. はじめに

古典的なゲーム理論の枠組みでは、有限繰り返し囚人のジレンマゲームにおいて、プレイヤーが後ろ向き帰納推論を行うことですべてのラウンドで非協力行動を選択する戦略が部分ゲーム完全均衡になる。一方、実験研究では、繰り返しゲームの後半に進むにつれて非協力行動が選択される頻度が高くなるものの、必ずしも協力的な行動選択が見られないわけではない。これを説明するための理論側の反応は、大きく次の 2 つに分類できるだろう。第一に、二者関係の記述をより精緻化するというものである。これには、あくまで合理的選択論の範疇で説明を試みるもの (Krep et al. 1982) と、通常のゲーム理論では想定されない外生的な要素を導入するもの (Montgomery 1998 など) とが考えられる。第二に、進化ゲーム理論によるアプローチである。Schuessler (1989) は決定論的な進化ダイナミクスを対象としているが、結局、より早いラウンドで非協力行動をとる 1 種類の戦略で集団が占められる様子しか描けていない。本研究は後者に位置づけられるものではあるが、有限集団における確率進化ダイナミクスという先行研究とは異なる観点から上述の問題に接近できる可能性を検討したい。

2. モデル

本研究では、有限集団における有限繰り返し囚人のジレンマゲームを対象とする。各ラウンドでプレイヤーは協力か非協力かのどちらかの行動を選択する。まず、繰り返しゲームの最終ラウンドから数えて i 番目のゲームから最終ラウンドまでを裏切る戦略を E_{i+1} と呼ぶ。よって、戦略名の添え字の値が小さいほど協力的な戦略であることを意味し、E_1 は AllC 戦略を、E_{m+1} は AllD 戦略を意味する。ここで、ラウンド数を m とすると、想定される戦略数は m+1 個になる。また、繰り返しゲームから協力の進化を考える先行研究の多く (Imhof et al. 2005; Antal et al. 2009; Kurokawa and Ihara 2009; Kurokawa et al. 2010) に倣い、本研究でも必要に応じて TFT 戦略を導入する。その場合、合計 m+2 個の戦略を想定することになる。

各ラウンドで協力者は、相手が協力行動を選択した場合は R の利得を、相手が非協力行

動を選択した場合は S の利得を得る。各ラウンドで非協力者は、相手が協力行動を選択した場合は T の利得を、相手が非協力行動を選択した場合は P の利得を得る。したがって、E_i 戦略が E_k 戦略とゲームした際に得られる総利得 a_{ik} は、
$i<k$ のとき

$$a_{ik} = R(m-k+1) + S(k-i) + P(i-1)$$

$i>k$ のとき

$$a_{ik} = R(m-k+1) + T(i-k) + P(k-1)$$

$i=k$ のとき

$$a_{ik} = R(m-k+1) + P(k-1)$$

となる。ただし以下では簡単のため、$R = b\text{-}c$、$S = \text{-}c$、$T = b$、$P = 0$ と仮定する。すなわち、b は相手から協力された際に得られる便益、c は自分が協力する際に生じるコストである。

プレイヤーはゲームの帰結に応じて戦略を変更する。ここでは先行研究 (Nowak et al. 2004; Fudenberg and Imhof 2006; Hauert et al. 2008; Traulsen and Hauert 2009) に倣い、以下で示される Moran 過程に従って戦略が更新されると仮定する。まず、各プレイヤーが集団中の全員とゲームをし、それに応じた適応度を計算する。ここでは戦略 i の適応度を $f_i \equiv \exp(-\delta(\mathbf{A}x)_i)$ と定義する。$\mathbf{A}$ は利得行列であり、x は戦略の頻度ベクトルである。次に、N 人の集団からランダムに 1 人が選ばれ、そのプレイヤーは適応度に比例して N 人中から選ばれたプレイヤーの戦略を採用する。戦略は 1-u の確率で引き継がれるが、u の確率で突然変異が生じ、n 戦略のうちランダムに戦略が選ばれる。$u > 0$ のとき、マルコフ連鎖が唯一の定常分布 π を持つ。

本稿では、有限集団における進化ダイナミクスを平衡頻度という指標で捉える。ここで平衡頻度について説明しておく。有限集団では戦略の頻度は以下で示されるような格子点によって表現される。

$$S_n^{(N)} \equiv \left\{ (x_1, \ldots, x_n) \in \mathbf{R}^n \mid \textstyle\sum_{i=1}^{n} x_i = 1, x_i \in \left\{ 0, \dfrac{1}{N}, \ldots, \dfrac{N-1}{N}, 1 \right\} \right\}$$

このとき、定常分布 π における戦略 x_i の平衡頻度とは、$\langle x_i \rangle_\delta \equiv \sum_{s \in S_n^{(N)}} x_i(s) \pi_\delta(s)$ のことである。ここで、添え字の $\delta \in [0,1]$ は選択の強さであり、これが 0 のときどの戦略も中立となり平衡頻度は $1/n$ (n は戦略数)となる。そこで以下では、$1/n$ を「中立レベルの平衡頻度」と呼び、平衡頻度が $1/n$ 以上の戦略は優位な戦略であると評価する。$n \times n$ 対称ゲームの平衡

頻度は既に Antal et al.(2009) によって導出されている（大槻 2009 も参照)。それによれば、

$$\langle x_i \rangle_\delta = \frac{1}{n} + \delta \frac{N(1-u)}{n(Nu+1)(Nu+2)}(L_i + NuH_i) + o(\delta)$$

となる。したがって、

$$\langle x_i \rangle_\delta > 1/n \Leftrightarrow L_i + NuH_i > 0$$

が成り立つ。ここで、

$$L_i \equiv \frac{1}{n}\sum_{k=1}^{n}(a_{ii} + a_{ik} - a_{ki} - a_{kk})$$

$$H_i \equiv \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}(a_{ik} - a_{jk})$$

である。

よって、Nu が小さいときは、

$$\langle x_i \rangle_\delta > 1/n \Leftrightarrow L_i > 0, \ \langle x_i \rangle_\delta > \langle x_j \rangle_\delta \Leftrightarrow L_i > L_j$$

が成り立ち、

Nu が大きいときは、

$$\langle x_i \rangle_\delta > 1/n \Leftrightarrow H_i > 0, \ \langle x_i \rangle_\delta > \langle x_j \rangle_\delta \Leftrightarrow H_i > H_j .$$

が成り立つ。

なお、分析の結果、E_i 戦略の平衡頻度は、ラウンド数があまりに小さくなければ i に対して単調増加であることが分かった。つまり、非協力的な戦略ほど平衡頻度は高いということである。以下の結果はそれを前提として各戦略の平衡頻度を算出したときに、全戦略中どれだけの戦略の平衡頻度が 1/n 以上になるか、という観点から集団内の協力傾向を評価することにしたい。そのため、Nu が小さいときは $L_i = 0$ となる i の値が、Nu が大きいときは $H_i = 0$ となる i の値が評価基準になる。この分析方法のイメージが図 1 である。

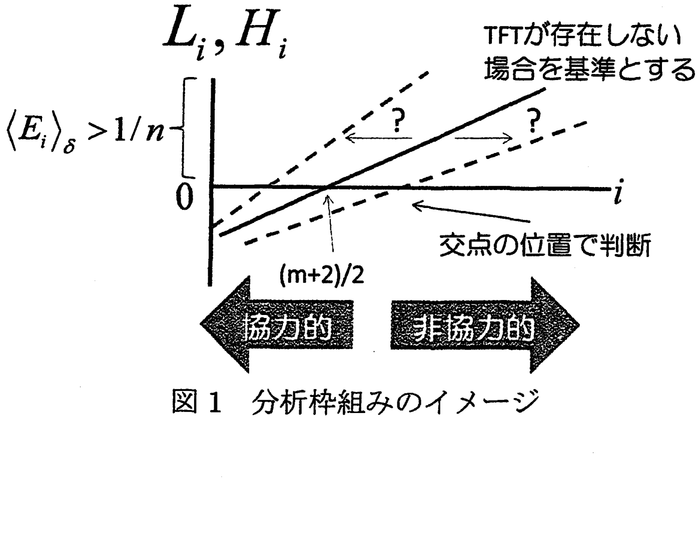

図 1　分析枠組みのイメージ

3. 結果

まずは TFT 戦略が存在しない場合の結果を示す。なお、以下の 3 つの結果については Nu の値に依存しない。

結果 1-1: どんな b、c、m に対しても、

結果 1-2: どんな b、c、m に対しても、平衡頻度が最大となるのは $i = m+1$ のとき。

結果 1-3: L_i、$H_i > 0$ となるのは $i > (m+2)/2$ のとき。

結果 1-1 は、AllC 戦略はいかなる条件下でも平衡頻度が中立レベルを上回れず、AllD 戦略はいかなる条件下でも平衡頻度が中立レベルを上回ることを意味し、結果 1-2 から、AllD 戦略は常に全戦略中最も平衡頻度が高い戦略であることを意味する。結果 1-3 は、$E_{(m+2)/2}$ 戦略よりも非協力的な戦略が中立レベル以上の平衡頻度を持つことを意味する。なお、$(m+2)/2$ は、繰り返しゲームの真ん中のラウンド番号である。

つづいて、結果 1 を参照点とし、TFT 戦略の存在が進化ダイナミクスにどのような影響を及ぼすのかを分析しよう。この場合、結果は Nu の大きさによって異なる。

結果2-1 (Nuが小さいとき):

$$L_i > 0 \text{ となるのは } i > \frac{2b - c(m^2 + 3m + 2)}{b - c(2m + 3)} \text{ のとき。}$$

結果2-2 (Nuが大きいとき):

$$H_i > 0 \text{ となるのは } i > \frac{cm^3 + (6c - b)m^2 + (11c - 9b)m + 4c - 6b}{2(m + 1)(cm + 3c - b)} \text{ のとき。}$$

上の 2 つの結果とも、右辺の値が高くなるほど、非協力的な戦略でなければ平衡頻度が中立レベルを上回れないことを意味する。ここで、数値例として、$b = 7$、$c = 1$、$m = 10$ の場合を考えよう(図 2 を参照)。まず、TFT 戦略が存在しないときは、結果 1-3 より、$i > 6$ のとき $\langle E_i \rangle_\delta > 1/n$ となる。つぎに、TFT が存在する場合であるが、Nu が小さいときは結果 2-1 より、$i > 7.375$ のとき、Nu が大きいときは結果 2-2 より、$i > 2.590$ のとき $\langle E_i \rangle_\delta > 1/n$ となる。これは、E_4 のような協力的な戦略は、TFT が存在し Nu が大きいときでないと中立レベル以上の平衡頻度になれないことを意味する。つまり、TFT が存在することで、より多くの協力的戦略の平衡頻度が中立レベルを上回るようになったのである。一方、TFT が存在し Nu が小さいときは、E_8、E_9、E_{10} のような非協力的な戦略以外は平衡頻度が全て中立レベルを下回ってしまっている。TFT が存在しない場合には E_8 の平衡頻度が中立レベルを

上回っていたのにも拘わらずである。つまり、TFT が存在することで、ゲームの早い段階で非協力行動をとる戦略しか平衡頻度が中立レベル以上にならなくなったのである。なお、さらなる数値計算の結果、このような関係が広いパラメータ領域で観察されることが分かった。

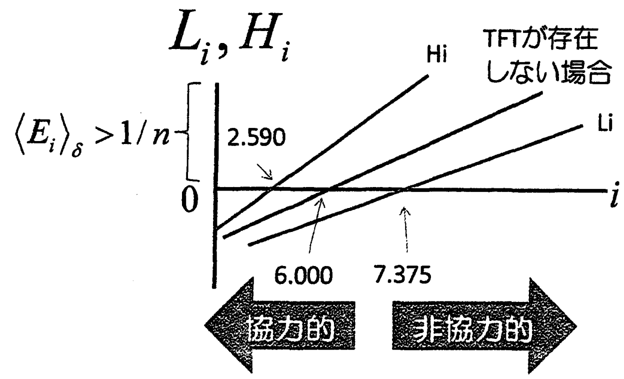

図２　数値例のイメージ
$b=7$、$c=1$、$m=10$ のとき。傾きは正確ではない。

4. 結論

結果として、*Nu* が小さい場合は、TFT が存在することで平衡頻度が中立レベルを上回る戦略が少なくなり、*Nu* が大きい場合には、TFT が存在するときの方が平衡頻度が中立レベルを上回る戦略が多くなる、つまり、より協力的な戦略でも中立レベル以上の平衡頻度を持つ、ということが広いパラメータ領域で成り立つことが分かった。

リプリケータダイナミクスに代表される侵入可能性によって戦略を評価する枠組みとは異なり、本稿で用いた平衡頻度という指標は、全ての戦略に対してその優位性を示す値を付与できるというメリットがある。そのため、Rand and Nowak (2012)も実践しているように、実験室実験などで得られた経験的データの戦略頻度の分布をよりよく説明できる可能性を秘めている。本研究は、Rand and Nowak (2012)とは評価方法が若干異なるものの、その一ケースといえるだろう。ただし本稿ではあくまでも、中立レベル以上の平衡頻度を持つ戦略の数を数えているだけであって、平衡頻度の値そのものを評価対象としているわけではない。この点は本稿の限界といえるであろう。また、今後はより高次な認知能力を要する戦略を導入した分析が望まれる。

参考文献

Antal, T., Traulsen, A., Ohtsuki, H., Tarnita, C.E., Nowak, M.A. 2009. "Mutation-selection equilibrium in games with multiple strategies." *Journal of Theoretical Biology* 258: 614-622.

Fudenberg, D., Imhof, L.A. 2006. "Imitation processes with small mutations." *Journal of Economic Theory* 131: 251-262.

Hauert, C., Traulsen, A., Brandt H., Nowak, M.A., Sigmund, K. 2008. "Public goods with punishment and abstaining in finite and infinite populations." *Biological Theory* 3: 114-122.

Imhof, L.A., Fudenberg, D., Nowak, M.A. 2005. "Evolutionary cycles of cooperation and defection." *Proceedings of the National Academy of Sciences of the United States of America* 102: 10797-10800.

Kreps, D.M., Milgrom, P., Roberts, J., Wilson, R. 1982. "Rational cooperation in the finitely repeated prisoners' dilemma." *Journal of Economic Theory* 27: 245-252.

Kurokawa, S., Ihara, Y. 2009. "Emergence of cooperation in public goods games." *Proceedings of the Royal Society B* 276: 1379-1384.

Kurokawa, S., Wakano, J. Y., Ihara, Y. 2010. "Generous cooperators can outperform non-generous cooperators when replacing a population of defectors." *Theoretical Population Biology* 77: 257-262.

Montgomery, J.D. 1998. "Toward a role-theoretic conception of embeddedness." *American Journal of Sociology* 104: 92-125.

Nowak, M. A., Sasaki, A., Taylor, C., Fudenberg, D. 2004. "Emergence of cooperation and evolutionary stability in finite populations." *Nature* 428: 646-650.

大槻久. 2009. 「*n*-戦略ゲームの mutation-selection balance 均衡」『京都大学数理解析研究所講究録』 1663: 118-123.

Rand, D.G., Nowak, M.A. 2012. "Evolutionary dynamics in finite populations can explain the full range of cooperative behaviors observed in the centipede game." *Journal of Theoretical Biology* 300: 212-221.

Schuessler, R. 1989. "The gradual decline of cooperation: Endgame effects in evolutionary game theory." *Theory and Decision* 26: 133-155.

Traulsen, A., Hauert, C. 2009. "Stochastic evolutionary game dynamics". in Schuster, H-G (eds). *Reviews of Nonlinear Dynamics and Complexity* Volume 2. Wiley-VCH: Weinheim.